SEQUENCE LISTING

<110>   Nilles, Matthew L.
        Matson, Jyl S.

<120>   YERSINIA SPECIES COMPOSITIONS

<130>   3128-6046US

<150>   US 60/444,076
<151>   2003-01-31

<160>   20

<170>   PatentIn version 3.2

<210>   1
<211>   87
<212>   PRT
<213>   Yersinia pestis

<220>
<223>   amino acid sequence of YscF

<220>
<223>   sequence can be found at MedLine accession number NP_857921.1

<400>   1

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu
1   5   10   15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys
            20   25   30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro
            35   40   45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile
    50   55   60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln
65   70   75   80

Gly Ile Leu Gln Lys Phe Pro
                85

<210> 2
<211> 264
<212> DNA
<213> Yersinia pestis

<220>
<223> nucliec acids encoding YscF

<220>
<223> sequence can be found at MedLine accession number NL_004839

<400> 2
atgagtaact tctctggatt tacgaaagga accgatatcg cagacttaga tgcggtggct   60

caaacgctca agaagccagc agacgatgca aacaaagcgg ttaatgactc gatagcagca   120

ttgaaagata agcctgacaa cccggcgcta cttgctgact acaacattc aattaataaa   180

tggtcggtaa tttacaatat aaactcaacc atagttcgta gcatgaaaga cttaatgcaa   240

ggcatcctac agaagttccc ataa   264

<210> 3
<211> 264
<212> DNA
<213> Yersinia enterocolitica

<220>
<223> nucleic acid encoding YscF from Y. enterocolitica

<220>
<223> sequence can be found at MedLine accession number NC_005017

<400> 3
atgagtaatt tctctgggtt tgcaaaagga accgatatca cagacttaga tgcggtggct   60

caaacgctca agaagccagc agacggcgca aacaaagcgg ttaatgactc gatagcagcg   120

ttgaaagaga cgcctgacaa cccggcgcta cttgctgact acaacattc aattaataaa   180

tggtcggtaa tttacaatat aagctcaacc atagttcgta gcatgaaaga cttaatgcaa   240

ggcatcctac agaagttccc ataa   264

<210> 4
<211> 87
<212> PRT

<213> Yersinia enterocolitica

<220>
<223> amino acid sequence of YscF from Y. enterocolitica

<220>
<223> sequence can be found at MedLine accession number NP_863538.1

<400> 4

Met Ser Asn Phe Ser Gly Phe Ala Lys Gly Thr Asp Ile Thr Asp Leu
1   5   10   15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys
            20   25   30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Thr Pro Asp Asn Pro
        35   40   45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile
    50   55   60

Tyr Asn Met Ser Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln
65   70   75   80

Gly Ile Leu Gln Lys Phe Pro
                85

<210> 5
<211> 80
<212> PRT
<213> Salmonella enterica

<220>
<223> amino acid sequence of PrgI

<220>
<223> sequence can be found at MedLine accession number CAD05980.1

<400> 5

Met Pro Thr Ser Trp Ser Gly Tyr Leu Asp Glu Val Ser Ala Lys Phe
1   5   10   15

3

```
Asp Lys Gly Val Asp Asn Leu Gln Thr Gln Val Thr Glu Ala Leu Asp
        20                  25                  30

Lys Leu Ala Ala Lys Pro Ser Asp Pro Ala Leu Leu Ala Ala Tyr Gln
        35                  40                  45

Ser Lys Leu Ser Glu Tyr Asn Leu Tyr Arg Asn Ala Gln Ser Asn Thr
    50                  55                  60

Val Lys Val Phe Lys Asp Ile Asp Ala Ala Ile Ile Gln Asn Phe Arg
65                  70                  75                  80
```

<210> 6
<211> 243
<212> DNA
<213> Salmonella enterica

<220>
<223> sequence can be found at MedLine accession number AL627276

<220>
<223> nucleic acid encoding PrgI

<400> 6
```
atgccaacat cttggtcagg ctatctggat gaagtttcag caaaatttga taagggcgtt      60

gataatctac aaacgcaggt aacagaggcg ctggataaat tagcagcaaa accctccgat     120

ccggcgctac tggcggcgta tcagagtaag ctctcggaat ataacttgta ccgtaacgcg     180

caatcgaaca cggtaaaagt ctttaaggat attgatgctg ccattattca gaacttccgt     240

taa                                                                    243
```

<210> 7
<211> 83
<212> PRT
<213> Shigella flexneri

<220>
<223> amino acid sequence of MxiH

<220>
<223> sequence can be found at MedLine accession number NP_858270.1

<400> 7

```
Met Ser Val Thr Val Pro Asn Asp Asp Trp Thr Leu Ser Ser Leu Ser
1               5                   10                  15


Glu Thr Phe Asp Asp Gly Thr Gln Thr Leu Gln Gly Glu Leu Thr Leu
            20                  25                  30


Ala Leu Asp Lys Leu Ala Lys Asn Pro Ser Asn Pro Gln Leu Leu Ala
            35                  40                  45


Glu Tyr Gln Ser Lys Leu Ser Glu Tyr Thr Leu Tyr Arg Asn Ala Gln
    50                  55                  60


Ser Asn Thr Val Lys Val Ile Lys Asp Val Asp Ala Ala Ile Ile Gln
65                  70                  75                  80


Asn Phe Arg
```

```
<210>   8
<211>   252
<212>   DNA
<213>   Shigella flexneri

<220>
<223>   nucleic acid encoding MxiH

<220>
<223>   sequence can be found at MedLine accession number Nc_004851

<400>   8
atgagtgtta cagtaccgaa tgatgattgg acattgagtt cattatctga aacttttgat      60

gatggaactc aaacattaca aggtgaacta acattggcac tagataaatt agctaaaaat     120

ccttcgaatc cacagttgct ggctgaatac caaagtaaat tatctgaata tacattatat     180

aggaacgcgc aatccaatac agtgaaagtg attaaggatg ttgatgctgc aattattcaa     240

aacttcagat aa                                                         252


<210>   9
<211>   73
<212>   PRT
```

<220>
<213> Escherichia coli

<220>
<223> amino acid sequence of EscF

<220>
<223> sequence can be found at MedLine accession number NP_312579.1

<400> 9

Met Asn Leu Ser Glu Ile Thr Gln Gln Met Gly Glu Val Gly Lys Thr
1   5   10   15

Leu Ser Asp Ser Val Pro Glu Leu Leu Asn Ser Thr Asp Leu Val Asn
            20   25   30

Asp Pro Glu Lys Met Leu Glu Leu Gln Phe Ala Val Gln Gln Tyr Ser
        35   40   45

Ala Tyr Val Asn Val Glu Ser Gly Met Leu Lys Thr Ile Lys Asp Leu
    50   55   60

Val Ser Thr Ile Ser Asn Arg Ser Phe
65   70

<210> 10
<211> 222
<212> DNA
<213> Escherichia coli

<220>
<223> nucleic acid encoding EscF

<220>
<223> sequence can be found at MedLine accession number NC_002695

<400> 10
atgaatttat ctgaaattac tcaacaaatg ggtgaagtag gtaaaacgct gagcgattct   60

gtgccagagt tacttaatag caccgatttg gttaatgacc ctgaaaaaat gttagagttg   120

cagtttgcgg ttcagcaata ttctgcttat gttaacgtag aaagtggaat gttgaaaacg   180

ataaaagatc tggtctcaac catttctaac cgtagttttt aa   222

6

<210> 11
<211> 240
<212> DNA
<213> Yersinia pestis

<220>
<223> nucleic acid encoding HT-YscF; an example of a recombinant YscF

<220>
<221> misc_feature
<222> (214)..(237)
<223> His-tag sequence

<400> 11
atgagtaact tctctggatt tacgaaagga accgatgatg caaacaaagc ggttaatgac   60

tcgatagcag cattgaaaga taagcctgac aacccggcgc tacttgctga cttacaacat   120

tcaattaata aatggtcggt aatttacaat ataaactcaa ccatagttcg tagcatgaaa   180

gacttaatgc aaggcatcct acagaagttc ccactcgagc accaccacca ccaccactga   240

<210> 12
<211> 95
<212> PRT
<213> Yersinia pestis

<220>
<223> amino acid sequence of HT-YscF; an example of a recombinant YscF

<220>
<221> misc_feature
<222> (88)..(95)
<223> His-tag

<400> 12

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu
1   5   10   15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys
            20   25   30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro
            35   40   45

```
Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile
    50                  55                  60


Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln
65                  70                  75                  80


Gly Ile Leu Gln Lys Phe Pro Leu Glu His His His His His
                85                  90                  95
```

```
<210>  13
<211>  5537
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  pMJ119

<220>
<221>  misc_feature
<222>  (563)..(1375)
<223>  kan encoded on complementary strand

<220>
<221>  misc_feature
<222>  (3518)..(4597)
<223>  LacI encoded on complementary strand

<220>
<221>  promoter
<222>  (4983)..(5000)
<223>  T7 promoter

<220>
<221>  misc_feature
<222>  (5001)..(5001)
<223>  T7 transcription start

<220>
<221>  misc_feature
<222>  (5072)..(5104)
<223>  T7 Tag

<220>
<221>  CDS
<222>  (5114)..(5401)
<223>  HT-YscF

<220>
```

<221> misc_feature
<222> (5381)..(5398)
<223> His-tag sequence

<220>
<221> terminator
<222> (5466)..(5512)
<223> T7 terminator

<400> 13

```
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg      60

cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc     120

ctttctcgcc acgttcgccg gctttccccg tcaagctcta atcgggggc tccctttagg      180

gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc     240

acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt     300

ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc     360

ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta     420

acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt     480

tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta     540

tccgctcatg aattaattct tagaaaaact catcgagcat caaatgaaac tgcaatttat     600

tcatatcagg attatcaata ccatatttt gaaaaagccg tttctgtaat gaaggagaaa      660

actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg attccgactc     720

gtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta tcaagtgaga     780

aatcaccatg agtgacgact gaatccggtg agaatggcaa aagtttatgc atttctttcc     840

agacttgttc aacaggccag ccattacgct cgtcatcaaa atcactcgca tcaaccaaac     900

cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac     960

aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat    1020

tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg gggatcgcag    1080

tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc ggaagaggca    1140

taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac    1200

ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat cgatagattg    1260
```

9

```
tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca   1320

tgttggaatt taatcgcggc ctagagcaag acgtttcccg ttgaatatgg ctcataacac   1380

cccttgtatt actgtttatg taagcagaca gttttattgt tcatgaccaa aatcccttaa   1440

cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga   1500

gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg   1560

gtggtttgtt tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc   1620

agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag   1680

aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc   1740

agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg   1800

cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac   1860

accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga   1920

aaggcggaca ggtatccggt aagcggcagg tcggaacag gagagcgcac gagggagctt   1980

ccaggggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag   2040

cgtcgatttt tgtgatgctc gtcaggggg cggagcctat ggaaaaacgc cagcaacgcg   2100

gcctttttac ggttcctggc cttttgctgg ccttttgctc acatgttctt tcctgcgtta   2160

tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc   2220

agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cctgatgcgg   2280

tattttctcc ttacgcatct gtgcggtatt cacaccgca tatatggtgc actctcagta   2340

caatctgctc tgatgccgca tagttaagcc agtatacact ccgctatcgc tacgtgactg   2400

ggtcatggct gcgccccgac acccgccaac acccgctgac gcgccctgac gggcttgtct   2460

gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag   2520

gttttcaccg tcatcaccga aacgcgcgag gcagctgcgg taaagctcat cagcgtggtc   2580

gtgaagcgat tcacagatgt ctgcctgttc atccgcgtcc agctcgttga gtttctccag   2640

aagcgttaat gtctggcttc tgataaagcg ggccatgtta agggcggttt tttcctgttt   2700

ggtcactgat gcctccgtgt aagggggatt tctgttcatg ggggtaatga taccgatgaa   2760
```

10

```
acgagagagg atgctcacga tacgggttac tgatgatgaa catgcccggt tactggaacg    2820

ttgtgagggt aaacaactgg cggtatggat gcggcgggac cagagaaaaa tcactcaggg    2880

tcaatgccag cgcttcgtta atacagatgt aggtgttcca cagggtagcc agcagcatcc    2940

tgcgatgcag atccggaaca taatggtgca gggcgctgac ttccgcgttt ccagacttta    3000

cgaaacacgg aaaccgaaga ccattcatgt tgttgctcag gtcgcagacg ttttgcagca    3060

gcagtcgctt cacgttcgct cgcgtatcgg tgattcattc tgctaaccag taaggcaacc    3120

ccgccagcct agccgggtcc tcaacgacag gagcacgatc atgcgcaccc gtggggccgc    3180

catgccggcg ataatggcct gcttctcgcc gaaacgtttg gtggcgggac cagtgacgaa    3240

ggcttgagcg agggcgtgca agattccgaa taccgcaagc gacaggccga tcatcgtcgc    3300

gctccagcga aagcggtcct cgccgaaaat gacccagagc gctgccggca cctgtcctac    3360

gagttgcatg ataaagaaga cagtcataag tgcggcgacg atagtcatgc cccgcgccca    3420

ccggaaggag ctgactgggt tgaaggctct caagggcatc ggtcgagatc ccggtgccta    3480

atgagtgagc taacttacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa    3540

cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat    3600

tgggcgccag ggtggttttt cttttcacca gtgagacggg caacagctga ttgcccttca    3660

ccgcctggcc ctgagagagt tgcagcaagc ggtccacgct ggtttgcccc agcaggcgaa    3720

aatcctgttt gatggtggtt aacggcggga tataacatga gctgtcttcg gtatcgtcgt    3780

atcccactac cgagatatcc gcaccaacgc gcagcccgga ctcggtaatg gcgcgcattg    3840

cgcccagcgc catctgatcg ttggcaacca gcatcgcagt gggaacgatg ccctcattca    3900

gcatttgcat ggtttgttga aaaccggaca tggcactcca gtcgccttcc cgttccgcta    3960

tcggctgaat ttgattgcga gtgagatatt tatgccagcc agccagacgc agacgcgccg    4020

agacagaact taatgggccc gctaacagcg cgatttgctg gtgacccaat gcgaccagat    4080

gctccacgcc cagtcgcgta ccgtcttcat gggagaaaat aatactgttg atgggtgtct    4140

ggtcagagac atcaagaaat aacgccggaa cattagtgca ggcagcttcc acagcaatgg    4200

catcctggtc atccagcgga tagttaatga tcagcccact gacgcgttgc gcgagaagat    4260

tgtgcaccgc cgctttacag gcttcgacgc cgcttcgttc taccatcgac accaccacgc    4320
```

11

```
tggcacccag ttgatcggcg cgagatttaa tcgccgcgac aatttgcgac ggcgcgtgca      4380

gggccagact ggaggtggca acgccaatca gcaacgactg tttgcccgcc agttgttgtg      4440

ccacgcggtt gggaatgtaa ttcagctccg ccatcgccgc ttccactttt tcccgcgttt      4500

tcgcagaaac gtggctggcc tggttcacca cgcgggaaac ggtctgataa gagacaccgg      4560

catactctgc gacatcgtat aacgttactg gtttcacatt caccaccctg aattgactct      4620

cttccgggcg ctatcatgcc ataccgcgaa aggttttgcg ccattcgatg gtgtccggga      4680

tctcgacgct ctcccttatg cgactcctgc attaggaagc agcccagtag taggttgagg      4740

ccgttgagca ccgccgccgc aaggaatggt gcatgcaagg agatggcgcc aacagtccc       4800

ccggccacgg ggcctgccac catacccacg ccgaaacaag cgctcatgag cccgaagtgg      4860

cgagcccgat cttccccatc ggtgatgtcg gcgatatagg cgccagcaac cgcacctgtg      4920

gcgccggtga tgccggccac gatgcgtccg gcgtagagga tcgagatctc gatcccgcga      4980

aattaatacg actcactata ggggaattgt gagcggataa caattcccct ctagaaataa      5040

ttttgtttaa ctttaagaag gagatataca tatggctagc atgactggtg gacagcaaat      5100

gggtcgggat ccg atg agt aac ttc tct gga ttt acg aaa gga acc gat       5149
            Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp
             1               5                  10

atc gca gac tta gat gcg gtg gct caa acg ctc aag aag cca gca gac        5197
Ile Ala Asp Leu Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp
         15                  20                  25

gat gca aac aaa gcg gtt aat gac tcg ata gca gca ttg aaa gat aag        5245
Asp Ala Asn Lys Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys
         30                  35                  40

cct gac aac ccg gcg cta ctt gct gac tta caa cat tca att aat aaa        5293
Pro Asp Asn Pro Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys
45                  50                  55                  60

tgg tcg gta att tac aat ata aac tca acc ata gtt cgt agc atg aaa        5341
Trp Ser Val Ile Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys
                     65                  70                  75

gac tta atg caa ggc atc cta cag aag ttc cca ctc gag cac cac cac        5389
Asp Leu Met Gln Gly Ile Leu Gln Lys Phe Pro Leu Glu His His His
                 80                  85                  90
```

12

```
cac cac cac tga gatccggctg ctaacaaagc ccgaaaggaa gctgagttgg      5441
His His His
        95


ctgctgccac cgctgagcaa taactagcat aaccccttgg ggcctctaaa cgggtcttga   5501

ggggtttttt gctgaaagga ggaactatat ccggat                            5537
```

```
<210>  14
<211>  5309
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  pET24b

<220>
<221>  terminator
<222>  (26)..(72)
<223>  T7 terminator

<220>
<221>  misc_feature
<222>  (140)..(157)
<223>  His-tag sequence

<220>
<221>  misc_feature
<222>  (206)..(238)
<223>  T7 Tag

<220>
<221>  misc_feature
<222>  (309)..(309)

<220>
<221>  misc_feature
<222>  (309)..(309)
<223>  T7 transcription start

<220>
<221>  promoter
<222>  (310)..(327)
<223>  T7 promoter

<220>
<221>  misc_feature
<222>  (713)..(716)
<223>  Alternate start codon "gtg"
```

<220>
<221> CDS
<222> (713)..(1792)
<223> lacI

<220>
<221> CDS
<222> (3935)..(4747)
<223> kan

<400> 14
atccggatat agttcctcct ttcagcaaaa aaccccctcaa gacccgttta gaggccccaa   60

ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc tttcgggctt   120

tgttagcagc cggatctcag tggtggtggt ggtggtgctc gagtgcggcc gcaagcttgt   180

cgacggagct cgaattcgga tcccgaccca tttgctgtcc accagtcatg ctagccatat   240

gtatatctcc ttcttaaagt taaacaaaat tatttctaga ggggaattgt tatccgctca   300

caattcccct atagtgagtc gtattaattt cgcgggatcg agatctcgat cctctacgcc   360

ggacgcatcg tggccggcat caccggcgcc acaggtgcgg ttgctggcgc ctatatcgcc   420

gacatcaccg atggggaaga tcgggctcgc cacttcgggc tcatgagcgc ttgtttcggc   480

gtgggtatgg tggcaggccc cgtggccggg ggactgttgg cgccatctc cttgcatgca   540

ccattccttg cggcggcggt gctcaacggc ctcaacctac tactgggctg cttcctaatg   600

caggagtcgc ataagggaga gcgtcgagat cccggacacc atcgaatggc gcaaaacctt   660

tcgcggtatg gcatgatagc gcccggaaga gagtcaattc agggtggtga at gtg aaa   718
                                                           Met Lys
                                                             1

cca gta acg tta tac gat gtc gca gag tat gcc ggt gtc tct tat cag   766
Pro Val Thr Leu Tyr Asp Val Ala Glu Tyr Ala Gly Val Ser Tyr Gln
         5   10   15

acc gtt tcc cgc gtg gtg aac cag gcc agc cac gtt tct gcg aaa acg   814
Thr Val Ser Arg Val Val Asn Gln Ala Ser His Val Ser Ala Lys Thr
     20   25   30

cgg gaa aaa gtg gaa gcg gcg atg gcg gag ctg aat tac att ccc aac   862
Arg Glu Lys Val Glu Ala Ala Met Ala Glu Leu Asn Tyr Ile Pro Asn
35   40   45   50

cgc gtg gca caa caa ctg gcg ggc aaa cag tcg ttg ctg att ggc gtt   910
Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser Leu Leu Ile Gly Val

14

gcc acc tcc agt ctg gcc ctg cac gcg ccg tcg caa att gtc gcg gcg   958
Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser Gln Ile Val Ala Ala
            70   75   80

att aaa tct cgc gcc gat caa ctg ggt gcc agc gtg gtg gtg tcg atg   1006
Ile Lys Ser Arg Ala Asp Gln Leu Gly Ala Ser Val Val Val Ser Met
            85   90   95

gta gaa cga agc ggc gtc gaa gcc tgt aaa gcg gcg gtg cac aat ctt   1054
Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala Ala Val His Asn Leu
            100   105   110

ctc gcg caa cgc gtc agt ggg ctg atc att aac tat ccg ctg gat gac   1102
Leu Ala Gln Arg Val Ser Gly Leu Ile Ile Asn Tyr Pro Leu Asp Asp
115   120   125   130

cag gat gcc att gct gtg gaa gct gcc tgc act aat gtt ccg gcg tta   1150
Gln Asp Ala Ile Ala Val Glu Ala Ala Cys Thr Asn Val Pro Ala Leu
                135   140   145

ttt ctt gat gtc tct gac cag aca ccc atc aac agt att att ttc tcc   1198
Phe Leu Asp Val Ser Asp Gln Thr Pro Ile Asn Ser Ile Ile Phe Ser
                150   155   160

cat gaa gac ggt acg cga ctg ggc gtg gag cat ctg gtc gca ttg ggt   1246
His Glu Asp Gly Thr Arg Leu Gly Val Glu His Leu Val Ala Leu Gly
            165   170   175

cac cag caa atc gcg ctg tta gcg ggc cca tta agt tct gtc tcg gcg   1294
His Gln Gln Ile Ala Leu Leu Ala Gly Pro Leu Ser Ser Val Ser Ala
            180   185   190

cgt ctg cgt ctg gct ggc tgg cat aaa tat ctc act cgc aat caa att   1342
Arg Leu Arg Leu Ala Gly Trp His Lys Tyr Leu Thr Arg Asn Gln Ile
195   200   205   210

cag ccg ata gcg gaa cgg gaa ggc gac tgg agt gcc atg tcc ggt ttt   1390
Gln Pro Ile Ala Glu Arg Glu Gly Asp Trp Ser Ala Met Ser Gly Phe
                215   220   225

caa caa acc atg caa atg ctg aat gag ggc atc gtt ccc act gcg atg   1438
Gln Gln Thr Met Gln Met Leu Asn Glu Gly Ile Val Pro Thr Ala Met
            230   235   240

ctg gtt gcc aac gat cag atg gcg ctg ggc gca atg cgc gcc att acc   1486
Leu Val Ala Asn Asp Gln Met Ala Leu Gly Ala Met Arg Ala Ile Thr
            245   250   255

gag tcc ggg ctg cgc gtt ggt gcg gat atc tcg gta gtg gga tac gac   1534

15

```
Glu Ser Gly Leu Arg Val Gly Ala Asp Ile Ser Val Val Gly Tyr Asp
    260                 265             270

gat acc gaa gac agc tca tgt tat atc ccg ccg tta acc acc atc aaa      1582
Asp Thr Glu Asp Ser Ser Cys Tyr Ile Pro Pro Leu Thr Thr Ile Lys
275                 280             285                 290

cag gat ttt cgc ctg ctg ggg caa acc agc gtg gac cgc ttg ctg caa      1630
Gln Asp Phe Arg Leu Leu Gly Gln Thr Ser Val Asp Arg Leu Leu Gln
                295             300             305

ctc tct cag ggc cag gcg gtg aag ggc aat cag ctg ttg ccc gtc tca      1678
Leu Ser Gln Gly Gln Ala Val Lys Gly Asn Gln Leu Leu Pro Val Ser
            310                 315             320

ctg gtg aaa aga aaa acc acc ctg gcg ccc aat acg caa acc gcc tct      1726
Leu Val Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Gln Thr Ala Ser
            325                 330             335

ccc cgc gcg ttg gcc gat tca tta atg cag ctg gca cga cag gtt tcc      1774
Pro Arg Ala Leu Ala Asp Ser Leu Met Gln Leu Ala Arg Gln Val Ser
    340                 345             350

cga ctg gaa agc ggg cag tgagcgcaac gcaattaatg taagttagct            1822
Arg Leu Glu Ser Gly Gln
355                 360


cactcattag gcaccgggat ctcgaccgat gcccttgaga gccttcaacc cagtcagctc   1882

cttccggtgg gcgcggggca tgactatcgt cgccgcactt atgactgtct tctttatcat   1942

gcaactcgta ggacaggtgc cggcagcgct ctgggtcatt ttcggcgagg accgctttcg   2002

ctggagcgcg acgatgatcg gcctgtcgct tgcggtattc ggaatcttgc acgccctcgc   2062

tcaagccttc gtcactggtc ccgccaccaa acgtttcggc gagaagcagg ccattatcgc   2122

cggcatggcg gccccacggg tgcgcatgat cgtgctcctg tcgttgagga cccggctagg   2182

ctggcggggt tgccttactg gttagcagaa tgaatcaccg atacgcgagc gaacgtgaag   2242

cgactgctgc tgcaaaacgt ctgcgacctg agcaacaaca tgaatggtct tcggtttccg   2302

tgtttcgtaa agtctggaaa cgcggaagtc agcgccctgc accattatgt tccggatctg   2362

catcgcagga tgctgctggc taccctgtgg aacacctaca tctgtattaa cgaagcgctg   2422

gcattgaccc tgagtgattt ttctctggtc ccgccgcatc cataccgcca gttgtttacc   2482

ctcacaacgt tccagtaacc gggcatgttc atcatcagta acccgtatcg tgagcatcct   2542
```

```
ctctcgtttc atcggtatca ttaccccat gaacagaaat cccccttaca cggaggcatc    2602

agtgaccaaa caggaaaaaa ccgcccttaa catggcccgc tttatcagaa gccagacatt    2662

aacgcttctg gagaaactca acgagctgga cgcggatgaa caggcagaca tctgtgaatc    2722

gcttcacgac cacgctgatg agctttaccg cagctgcctc gcgcgtttcg gtgatgacgg    2782

tgaaaacctc tgacacatgc agctcccgga cacggtcaca gcttgtctgt aagcggatgc    2842

cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggcgcagc    2902

catgacccag tcacgtagcg atagcggagt gtatactggc ttaactatgc ggcatcagag    2962

cagattgtac tgagagtgca ccatatatgc ggtgtgaaat accgcacaga tgcgtaagga    3022

gaaaataccg catcaggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg    3082

ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta ccacagaat     3142

cagggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta    3202

aaaaggccgc gttgctggcg ttttttccata ggctccgccc ccctgacgag catcacaaaa    3262

atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc    3322

cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt    3382

ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca    3442

gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccccc gttcagcccg    3502

accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat    3562

cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta    3622

cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct    3682

gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac    3742

aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa    3802

aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa    3862

actcacgtta agggattttg gtcatgaaca ataaaactgt ctgcttacat aaacagtaat    3922

acaagggggtg tt atg agc cat att caa cgg gaa acg tct tgc tct agg ccg    3973
                Met Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg Pro
                           365                     370

cga tta aat tcc aac atg gat gct gat tta tat ggg tat aaa tgg gct    4021
```

17

```
    Arg Leu Asn Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala
        375                 380             385

    cgc gat aat gtc ggg caa tca ggt gcg aca atc tat cga ttg tat ggg    4069
    Arg Asp Asn Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly
    390                 395             400                 405

    aag ccc gat gcg cca gag ttg ttt ctg aaa cat ggc aaa ggt agc gtt    4117
    Lys Pro Asp Ala Pro Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val
                        410             415             420

    gcc aat gat gtt aca gat gag atg gtc aga cta aac tgg ctg acg gaa    4165
    Ala Asn Asp Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu
                    425             430             435

    ttt atg cct ctt ccg acc atc aag cat ttt atc cgt act cct gat gat    4213
    Phe Met Pro Leu Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp
                440             445             450

    gca tgg tta ctc acc act gcg atc ccc ggg aaa aca gca ttc cag gta    4261
    Ala Trp Leu Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val
            455             460             465

    tta gaa gaa tat cct gat tca ggt gaa aat att gtt gat gcg ctg gca    4309
    Leu Glu Glu Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala
    470                 475             480                 485

    gtg ttc ctg cgc cgg ttg cat tcg att cct gtt tgt aat tgt cct ttt    4357
    Val Phe Leu Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe
                    490             495             500

    aac agc gat cgc gta ttt cgt ctc gct cag gcg caa tca cga atg aat    4405
    Asn Ser Asp Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn
                    505             510             515

    aac ggt ttg gtt gat gcg agt gat ttt gat gac gag cgt aat ggc tgg    4453
    Asn Gly Leu Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp
                520             525             530

    cct gtt gaa caa gtc tgg aaa gaa atg cat aaa ctt ttg cca ttc tca    4501
    Pro Val Glu Gln Val Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser
                535             540             545

    ccg gat tca gtc gtc act cat ggt gat ttc tca ctt gat aac ctt att    4549
    Pro Asp Ser Val Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile
    550                 555             560                 565

    ttt gac gag ggg aaa tta ata ggt tgt att gat gtt gga cga gtc gga    4597
    Phe Asp Glu Gly Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly
                        570             575             580
```

18

```
atc gca gac cga tac cag gat ctt gcc atc cta tgg aac tgc ctc ggt      4645
Ile Ala Asp Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly
        585                 590                 595

gag ttt tct cct tca tta cag aaa cgg ctt ttt caa aaa tat ggt att      4693
Glu Phe Ser Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile
        600                 605                 610

gat aat cct gat atg aat aaa ttg cag ttt cat ttg atg ctc gat gag      4741
Asp Asn Pro Asp Met Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu
        615                 620                 625

ttt ttc taagaattaa ttcatgagcg gatacatatt tgaatgtatt tagaaaaata      4797
Phe Phe
630

aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgaaatt gtaaacgtta      4857

atatttgtt aaaattcgcg ttaaattttt gttaaatcag ctcatttttt aaccaatagg      4917

ccgaaatcgg caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg      4977

ttccagtttg gaacaagagt ccactattaa agaacgtgga ctccaacgtc aaagggcgaa      5037

aaaccgtcta tcagggcgat ggcccactac gtgaaccatc accctaatca agttttttgg      5097

ggtcgaggtg ccgtaaagca ctaaatcgga accctaaagg gagcccccga tttagagctt      5157

gacggggaaa gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa ggagcgggcg      5217

ctagggcgct ggcaagtgta gcggtcacgc tgcgcgtaac caccacccc gccgcgctta      5277

atgcgccgct acagggcgcg tcccattcgc ca                                    5309
```

<210> 15
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> primer, HT-YscF Start

<400> 15
ccggatccga tgagtaactt ctctggattt   30

<210> 16
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> primer, HT-YscF Stop

<400> 16
ccgctcgagt gggaacttct gtaggatgcc   30

<210> 17
<211> 87
<212> PRT
<213> Yersinia pestis KIM5

<220>
<223> YscF amino acid seuquence from Y. pestis KIM5

<220>
<223> sequence can be found at MedLine accession number NC_004839

<400> 17

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu
1   5   10   15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys
            20   25   30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro
            35   40   45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile
        50   55   60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln
65   70   75   80

Gly Ile Leu Gln Lys Phe Pro
                85

<210> 18
<211> 87
<212> PRT
<213> Yersinia pestis CO92

<220>

<223> YscF amino acid sequence from Y. pestis CO92

<220>
<223> sequence can be found at MedLine accession number NC_003131

<400> 18

```
Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu
1               5                   10                  15


Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys
            20                  25                  30


Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro
        35                  40                  45


Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile
    50                  55                  60


Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln
65                  70                  75                  80


Gly Ile Leu Gln Lys Phe Pro
                85
```

<210> 19
<211> 87
<212> PRT
<213> Artificial Sequence

<220>
<223> Consensus sequence for YscF

<400> 19

```
Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu
1               5                   10                  15


Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys
            20                  25                  30


Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro
        35                  40                  45
```

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile
    50   55   60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln
65   70   75   80

Gly Ile Leu Gln Lys Phe Pro
                85

<210>   20
<211>   4884
<212>   DNA
<213>   Artificial Sequence

<220>
<223>   pBAD18-YscF

<220>
<221>   misc_feature
<222>   (96)..(974)
<223>   araC

<220>
<221>   misc_feature
<222>   (1003)..(1020)
<223>   operator 02

<220>
<221>   promoter
<222>   (1125)..(1153)
<223>   Pc promoter

<220>
<221>   misc_feature
<222>   (1161)..(1182)
<223>   operator 01

<220>
<221>   misc_feature
<222>   (1204)..(1217)
<223>   CAP site

<220>
<221>   misc_feature
<222>   (1213)..(1251)

<223> operator I2 + I1

<220>
<221> promoter
<222> (1250)..(1277)
<223> PBAD promoter

<220>
<221> CDS
<222> (1323)..(1586)
<223> YscF

<220>
<221> misc_feature
<222> (1634)..(2059)
<223> rrnB

<220>
<221> terminator
<222> (1634)..(2059)
<223> rrnB

<220>
<221> promoter
<222> (2114)..(2120)
<223> bla P3 promoter

<220>
<221> misc_feature
<222> (2152)..(3015)
<223> bla

<220>
<221> rep_origin
<222> (3051)..(3509)
<223> M13 origin

<220>
<221> rep_origin
<222> (3515)..(4212)
<223> pBR322 origin

<400> 20
atcgatgcat aatgtgcctg tcaaatggac gaagcaggga ttctgcaaac cctatgctac   60

tccgtcaagc cgtcaattgt ctgattcgtt accaattatg acaacttgac ggctacatca   120

ttcactttttt cttcacaacc ggcacggaac tcgctcgggc tggccccggt gcattttttta   180

aatacccgcg agaaatagag ttgatcgtca aaaccaacat tgcgaccgac ggtggcgata   240

```
ggcatccggg tggtgctcaa aagcagcttc gcctggctga tacgttggtc ctcgcgccag      300

cttaagacgc taatccctaa ctgctggcgg aaaagatgtg acagacgcga cggcgacaag      360

caaacatgct gtgcgacgct ggcgatatca aaattgctgt ctgccaggtg atcgctgatg      420

tactgacaag cctcgcgtac ccgattatcc atcggtggat ggagcgactc gttaatcgct      480

tccatgcgcc gcagtaacaa ttgctcaagc agatttatcg ccagcagctc cgaatagcgc      540

ccttcccctt gcccggcgtt aatgatttgc ccaaacaggt cgctgaaatg cggctggtgc      600

gcttcatccg ggcgaaagaa ccccgtattg gcaaatattg acggccagtt aagccattca      660

tgccagtagg cgcgcggacg aaagtaaacc cactggtgat accattcgcg agcctccgga      720

tgacgaccgt agtgatgaat ctctcctggc gggaacagca aaatatcacc cggtcggcaa      780

acaaattctc gtccctgatt tttcaccacc ccctgaccgc gaatggtgag attgagaata      840

taacctttca ttcccagcgg tcggtcgata aaaaaatcga dataaccgtt ggcctcaatc      900

ggcgttaaac ccgccaccag atgggcatta aacgagtatc ccggcagcag gggatcattt      960

tgcgcttcag ccatactttt catactcccg ccattcagag aagaaaccaa ttgtccatat     1020

tgcatcagac attgccgtca ctgcgtcttt tactggctct tctcgctaac caaaccggta     1080

accccgctta ttaaaagcat tctgtaacaa agcgggacca aagccatgac aaaaacgcgt     1140

aacaaaagtg tctataatca cggcagaaaa gtccacattg attatttgca cggcgtcaca     1200

ctttgctatg ccatagcatt tttatccata agattagcgg atcctacctg acgctttta      1260

tcgcaactct ctactgtttc tccatacccg ttttttgggg ctagcgaatt caggaggaaa     1320

cg atg agt aac ttc tct gga ttt acg aaa gga acc gat atc gca gac        1367
   Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp
   1                5                  10                  15

tta gat gcg gtg gct caa acg ctc aag aag cca gca gac gat gca aac      1415
Leu Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn
             20                  25                  30

aaa gcg gtt aat gac tcg ata gca gca ttg aaa gat aag cct gac aac      1463
Lys Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn
             35                  40                  45

ccg gcg cta ctt gct gac tta caa cat tca att aat aaa tgg tcg gta      1511
Pro Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val
          50                  55                  60
```

24

```
att tac aat ata aac tca acc ata gtt cgt agc atg aaa gac tta atg      1559
Ile Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met
     65              70              75

caa ggc atc cta cag aag ttc cca taa ggatcccgcg gggatcctct          1606
Gln Gly Ile Leu Gln Lys Phe Pro
80              85

agagtcgacc tgcaggcatg caagcttggc tgttttggcg datgagagaa gattttcagc   1666

ctgatacaga ttaaatcaga acgcagaagc ggtctgataa aacagaattt gcctggcggc   1726

agtagcgcgg tggtcccacc tgaccccatg ccgaactcag aagtgaaacg ccgtagcgcc   1786

gatggtagtg tggggtctcc ccatgcgaga gtagggaact gccaggcatc aaataaaacg   1846

aaaggctcag tcgaaagact gggcctttcg ttttatctgt tgtttgtcgg tgaacgctct   1906

cctgagtagg acaaatccgc cgggagcgga tttgaacgtt gcgaagcaac ggcccggagg   1966

gtggcgggca ggacgcccgc cataaactgc caggcatcaa attaagcaga aggccatcct   2026

gacggatggc ctttttgcgt ttctacaaac tcttttgttt atttttctaa atacattcaa   2086

atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga   2146

agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg gcattttgcc   2206

ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg   2266

gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc   2326

gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat   2386

tatcccgtgt tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg   2446

acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag   2506

aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa   2566

cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat catgtaactc   2626

gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca   2686

cgatgcctgc agcaatggca acaacgttgc gcaaactatt aactggcgaa ctacttactc   2746

tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc   2806

tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg   2866
```

25

```
ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta    2926

tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag    2986

gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga    3046

ttgatttacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc    3106

gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt    3166

ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc    3226

cgatttagtg ctttacggca cctcgacccc aaaaaacttg atttgggtga tggttcacgt    3286

agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt    3346

aatagtggac tcttgttcca aacttgaaca acactcaacc ctatctcggg ctattctttt    3406

gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa    3466

aaatttaacg cgaattttaa caaaatatta acgtttacaa tttaaaagga tctaggtgaa    3526

gatccttttt gataatctca tgaccaaaat cccttaacgt gagttttcgt tccactgagc    3586

gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat    3646

ctgctgcttg caaacaaaaa aaccaccgct accagcggtg gtttgtttgc cggatcaaga    3706

gctaccaact ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt    3766

ccttctagtg tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata    3826

cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac    3886

cgggttggac tcaagacgat agttaccgga taaggcgcag cggtcgggct gaacgggggg    3946

ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg    4006

tgagctatga gaaagcgcca cgcttcccga agggagaaag gcggacaggt atccggtaag    4066

cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct    4126

ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgattttgt gatgctcgtc    4186

agggggggcgg agcctatgga aaaacgccag caacgcggcc ttttacggt tcctggcctt    4246

ttgctggcct tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg    4306

tattaccgcc tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga    4366

gtcagtgagc gaggaagcgg aagagcgcct gatgcggtat tttctcctta cgcatctgtg    4426
```

26

```
cggtatttca caccgcatat ggtgcactct cagtacaatc tgctctgatg ccgcatagtt    4486

aagccagtat acactccgct atcgctacgt gactgggtca tggctgcgcc ccgacacccg    4546

ccaacacccg ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc ttacagacaa    4606

gctgtgaccg tctccgggag ctgcatgtgt cagaggtttt caccgtcatc accgaaacgc    4666

gcgaggcagc aaggagatgg cgcccaacag tcccccggcc acggggcctg ccaccatacc    4726

cacgccgaaa caagcgctca tgagcccgaa gtggcgagcc cgatcttccc catcggtgat    4786

gtcggcgata taggcgccag caaccgcacc tgtggcgccg gtgatgccgg ccacgatgcg    4846

tccggcgtag aggatctgct catgtttgac agcttatc                            4884
```